Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD, Oct 1-5, 2007.

NAÏVE REALISM IN TERRAIN APPRECIATION

Harvey S. Smallman, Maia B. Cook, Daniel I. Manes,
Pacific Science and Engineering Group, Inc., San Diego, CA
and
Michael B. Cowen
Space and Naval Warfare Systems Center, San Diego, CA

Previously, we have shown that shaded perspective view ("3-D") displays are better for understanding the shape and rough layout of terrain than conventional 2-D views. We have coined the term *Naïve Realism* for users' misplaced, blanket faith in these 3-D displays (Smallman & St. John, 2005). There are hints in the individual difference literature that those of low spatial ability may be particularly prone to Naïve Realism. Here, we integrate these notions to test several theoretical predictions and to develop a new terrain simplification concept. Thirty-three participants had their spatial ability and problem-solving style measured. Then participants predicted which displays would, and then did, best support them in performing a task of threading a concealed route through realistic terrain. Depth relief (shading vs. topographic lines), viewing angle (90° vs. 45°) and terrain fidelity (high/unfiltered sharp vs. low/spatially smoothed) were all varied. Of the eight display combinations, Naïve Realism correctly predicted the greatest preference for the highest fidelity, realistic 3-D view (sharp, shaded, 45°). Yet the routing task was best performed with lower fidelity views. Spatially filtering terrain unmasks canyons and other gross terrain features, enabling them to pop-out more easily. Individuals of high spatial ability had better task performance and better calibrated their post-task display preferences, suggesting they are generally more savvy about the ways that display format affects their performance.

INTRODUCTION

Many operational tasks require understanding the shape and layout of three-dimensional (3-D) scenes. For example, users may need to understand lines of sight for reconnaissance, locate promising avenues for routing strikes and for arranging communications, and so on. In order to successfully perform these tasks, users are often said to require "terrain appreciation."

Classically, users have had to get their terrain appreciation from top-down, 2-D views of terrain. Depth relief on these 2-D views has most commonly been given by topographic, or "topo," lines (so-called "topo maps"). It has long been appreciated, though, that topo maps are challenging to rapidly and accurately mentally reconstruct in three dimensions (e.g., Pick, Heinrichs, Montello, et al., 1995). Accordingly, display designers have created 3-D perspective views of terrain, or so-called "3-D maps" (Jenks & Brown, 1966) which appear to dramatically convey three-dimensional structure and shape needed for terrain appreciation.

However, we have shown that 3-D displays are only superior for a subset of terrain appreciation tasks such as judging the rough layout of a scene. When any precise local judgment is required, 2-D displays are superior (St. John, Cowen, Smallman, & Oonk, 2001a). Indeed, terrain appreciation is not a monolithic concept. Routing a chain of communication antennas through terrain, for example, involves at least two subtasks that pose different requirements and are best served by different display configurations. Initially determining promising avenues through terrain involves appreciating the coarse 3-D scene layout and structure, and it was found to be better served by a shaded 3-D view of the terrain. Subsequently checking clearances and lines of sight involves appreciating specific, local scene characteristics, and it was found to be better served by a 2-D topo map of the terrain (St. John, Smallman, Banks, & Cowen, 2001b).

Here, we explore what aspect of 3-D displays makes them superior for shape understanding? Is it the shallow viewing angle or the shaded relief, or both?

Another factor that plays into the problem are users' beliefs about what displays may be best for which tasks. Smallman and St. John (2005) have coined the term *Naïve Realism* to label users' misplaced, blanket faith in realistic 3-D displays with which they then under-perform. One aspect of Naïve Realism that we explore here is a desire for full fidelity. Users may believe that the highest fidelity, realistic representation will always serve them best. To extract gross scene layout, though, it is possible that lowering the fidelity of terrain by simplifying it may increase performance. By spatially-filtering ("smoothing") the terrain, fine details may be obliterated that may otherwise mask the detection of gross scene features needed to appreciate layout - we can help users see the forest by getting rid of the trees. Note how canyons "pop" better from the smoothed terrain, see Figure 1, right. Less can be more. However, Naïve Realism predicts that users will still opt for full fidelity in this case.

Until now, we have generally ignored data

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **Naive Realism in Terrain Appreciation**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Pacific Science & Engineering Group,9180 Brown Deer Road,San Diego,CA,92121**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

variability in our 3-D experiments. However, examining variability *per se* may have explanatory power that can help us refine the Naïve Realism concept. Individuals differ in their spatial ability and style of problem-solving and both abilities have been shown to affect performance with visual displays. For example, Savage, Wiebe & Devine (2004) showed that those of higher spatial ability performed better on shape understanding tasks. In addition, those of lower spatial ability whose style is to solve problems visually (so-called "low-spatial visualizers") have difficulty interpreting abstract spatial representations such as graphs and constructing problem representations that extract only the relevant information from the problem (Hegarty & Kozhevnikov, 1999). Is it the case that low-spatial visualizers are particularly poor at predicting which displays to use for different tasks?

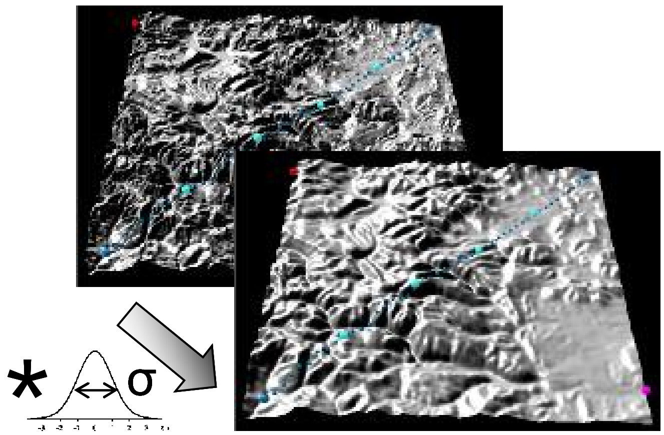

Figure 1. How terrain fidelity was modified with Gaussian low-pass spatial filtering from high/sharp (left) to low/smoothed (right). Shaded 45º scene views are shown.

Here, we integrate these previously separate lines of work on performance and preference, intuition and individual differences to both test and refine the concept of Naïve Realism and to develop a new, terrain appreciation display concept in the domain of a quasi-operational route planning task.

EXPERIMENT

Method

Participants. Thirty-three college students or graduates (12 male, 21 female) were recruited from www.CraigsList.com and were paid $30 for their participation. They had a mean age of 32.2 years (range 18-61 yrs).

Design. Three independent variables were manipulated in a 2 x 2 x 2 fully repeated measures design. The variables were viewing angle (45º vs. 90º), depth relief format (shading vs. topo), and terrain fidelity (high/unfiltered sharp vs. low/spatially smoothed). For each condition, a scene view was created that represented the intersection of the three variables, making eight scene views in total. The routing was broken down into four phases, with route laying occurring in Phases 1 & 3, and route elevation judgments occurring in Phases 2 & 4. All 33 participants performed these four phases on the eight scene views. Each scene view was of a different swath of terrain so that participants couldn't carry over terrain knowledge from view to view. The eight terrain swaths were always presented in the same order, but the order of the scene views, and therefore the scene view-terrain pairing, was counterbalanced across participants.

Stimuli. Each scene view consisted of a rendering of an approx. 4 by 5 mile swath of terrain. Start and finish route locations were indicated by large dark blue dots at opposite corners of the terrain (start at front left, and finish at back right), see Fig 1. An initial route between the start and end locations was shown by four, equally spaced waypoints, shown by large light blue dots. Smaller blue dots defined the segments between the waypoints. Participants could select and drag the waypoints to different positions on the terrain to create and indicate their preferred route through the terrain.

Eight terrain swaths were created from a selection of U.S. Geological digital elevation models (DEMs) of East San Diego county that we used in earlier research (see St. John et al., 2001b). Terrain difficulty was roughly equated by grading each candidate swath against a list of six criteria and then by normalizing the chosen swaths to possess the same altitude range.

The scenes were shown in two relief formats that were created with similar "texture draping" procedures. This draping procedure enabled both formats to be rendered from different viewing angles. Shaded relief was created by draping a grey matte texture and then lighting it from the conventional NW direction, see Fig 1. Topo relief was created by draping a white texture over the terrain mesh and then adding appropriate color-coded contour lines for equal altitude increments. There was no shading in the topo format. A color legend and scale were added to the side of the topo view. In both reliefs, colored dots were shown at the locations of the highest and lowest elevations on the scenes in the appropriate color from the topo legend to facilitate scene interpretation.

Terrain fidelity was manipulated by spatially filtering the terrain DEMs before meshing and rendering them. Custom software convolved the DEMs with a Gaussian low-pass spatial filter of space constant (σ) 3 pixels. This extent of smoothing was determined with pilot work that experimented with different σ values and settled on one that appeared appropriate for the complexity of the specific terrain swaths we were working with. Finally, the scene views were rendered in perspective from either 45º or 90º viewing angles using standard camera geometry.

Procedure. After informed consent, participants were administered the classic Vandenberg Mental Rotation Test (MRT) of spatial ability, and the Cognitive Style test of Kozhevnikov, Hegarty, & Mayer's (2002). The Style test classifies each individual as a verbalizer or a visualizer and then further classifies the latter as either high-spatial or low-spatial from their MRT score.

Participants were then asked to role-play an Army surveyor laying concealed routes through unfamiliar terrain. Good routes were defined to them as ones that were masked from as much of the surrounding terrain as possible. Good routes might hug narrow canyons while bad routes might run along ridges and over hilltops. Example good and bad route routes were shown. The different display formats and the HCI for waypoint interaction were explained in detail.

The procedures for the four phases of the routing experiment were then explained. The phases are laid out below, each followed by the dependent variables (DVs) recorded at that phase.

Phase 1 (Initial Route Laying): A straight line route from start to finish was laid out across the terrain, defined by four equally spaced waypoints. These waypoints had to be adjusted, as quickly as possible, to lay out a rough concealed route through the terrain.

DV: Initial route laying time.

Phase 2 (Initial Route Altitude): Participants reconstructed the altitude profile of the four waypoints in their initial route on a profile view by selecting and sliding each up and down a vertical slider. The highest and lowest altitudes from the map were shown on the slider, as were the start and endpoints of the route.

DV: Elevation inaccuracy (%)

Phase 3 (Final Route Adjustment): The initial route from Phase 1 was now defined by 14 waypoints that had to be carefully adjusted to lay out a final, maximally concealed route through the terrain. DV: Relative (%) improvement in route masking from initial to final route.

We developed a new terrain appreciation metric to assess route concealment. Masking was defined as the mean length of an array of lines of sight (LoS) shot out orthogonally all along the route (yellow lines in Fig 2) until they hit the furthest point the route could be seen from. Long LoS lengths meant a worse masked route had been laid. Short LoS lengths meant a better masked route had been laid, see Figure 2.

Phase 4 (Final Route Altitude): Participants reconstructed the altitude profile of the 14 waypoints, just as they had done with the four waypoints in Phase 2. DV: Elevation inaccuracy (%).

After instructions on the four phases, participants were shown a piece of practice terrain in all eight possible scene views and were asked to predict which would best support each phase of the experiment. Participants then practiced on that piece of terrain in their first assigned format and then completed all eight experimental views. No feedback was given on task performance during the experiment. When they were finished, participants were asked which view they thought supported their best performance for each phase. They were also asked to choose the one display that they would prefer to work with if required in daily use.

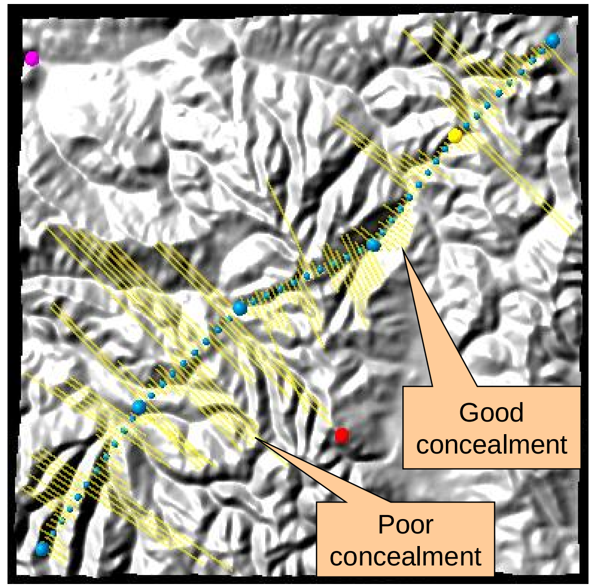

Figure 2. New line of sight length metric of concealment.

It took about 90 minutes to complete the entire procedure, including psychometric testing and practice.

Results

Only a subset of the entire dataset and analyses can be reported in this format. We hit the main points below.

Psychometric testing: Mean MRT score was 28.8 ± 3.2 (range 0 - 74), 34.8 for males and 25.4 for females. Of the 33 participants, there were 8 verbalizers, 10 high-spatial visualizers, and 15 low-spatial visualizers.

Intuitions: Participants predicted they would perform the masking best with a sharp, shaded, 45º (χ^2s, $p < .05$ for each attribute), see Figure 3, below. Figure 3 shows the overwhelming (75%) preference for sharp, shaded displays. The top four chosen displays were all sharp.

For Phases 2 & 4, only after the study did Participants significantly predict they would perform the altitude reconstruction better with topo vs. shaded displays (χ^2 $p < .05$).

Performance - Phase 1 route laying RTs: There was a main effect of relief format ($F(1, 25) = 9.7$, $p < .01$). Participants chose and laid the initial route quicker with shaded ($M = 39.3$ sec) than topo relief ($M = 46.0$ sec).

There were no main effects of viewing angle or fidelity. There was a format by detail interaction (F(1,25)=6.5, p<.05) with shaded views only faster than topo when sharp. Smoothing speeded route laying for topo views by reducing and smoothing the number of lines to interpret.

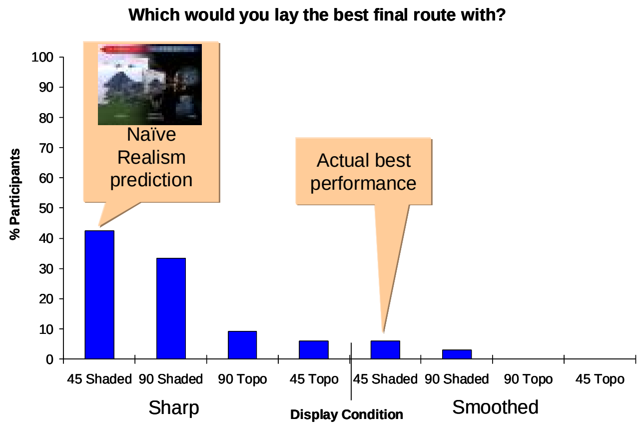

Figure 3. Intuitions about best display for final routing in descending order of preference. Arrows show Naïve Realism prediction (left) and actual best performing display (right).

Performance - Phases 2 & 4 altitude reconstruction: In Phase 2, there was a main effect of relief format ($F(1, 25) = 65.17$, $p < .001$). Participants reconstructed the altitudes of waypoints more accurately with topo relief than shaded relief (13.1% vs. 21.4% unsigned error). There were no main effects of viewing angle or fidelity. The same performance pattern also obtained in Phase 4.

Performance - Phase 3 masking: There was a main effect of relief ($F(1, 25) = 12.4$, $p < .01$), with shading promoting more improvement in masking than topo (10.0% vs. 1.3%). Masking was improved by shifting the viewing angle to 45˚ for views with shaded relief (format by viewing angle interaction of $F(1, 25) = 4.7$, $p < .05$). Finally, there was a 3 way interaction of fidelity by format by viewing angle ($F(1, 25) = 6.9$, $p < .05$). Smoothing improved masking at 45˚ only for shaded displays: the opposite result held at 90˚, see Figure 4. The best final routing performance came from the low fidelity (smoothed), shaded, 45º view ($p < .05$ by Tukey/Kramer post hoc tests). It performed twice as well as the sharp, shaded 45º view that participants intuited would perform best, see Fig 4.

Spatial ability & performance: For Phase 3, those with higher MRTs had significantly improved masking ($r(32) = .4$, $p < .05$). For Phases 2 & 4, those with higher MRTs had significantly reduced elevation error ($r(32) = -.4$, $p < .05$).

Intuitions & cognitive style: When predicting which displays they would perform best with before and after the Phase 3 route task, a trend emerged when Participants were broken down by cognitive style. Before/after proportions favoring sharp over smoothed were largely unchanged for the verbalizers and low-spatial visualizers (.91 vs. .87) whereas high-spatial visualizers pre/post proportions dropped from .90 to .60.

Preference: There was a significant preference to look at sharp, realistically shaded displays (χ^2 $p < .05$ for both fidelity and relief attributes). Participants did not differ significantly in their preference for a 45º vs. 90º viewing angle.

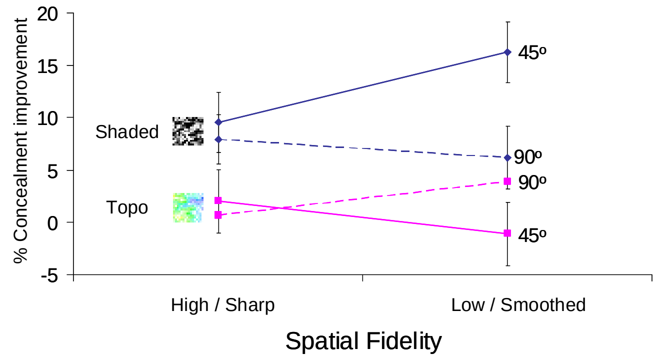

Figure 4. Improvement in route concealment afforded by different display formats.

DISCUSSION

In the *Science* paper that gave birth to 3-D terrain maps forty years ago, Jenks and Brown (1966) noted, presciently, that *"the cartographer…must choose between realism and practicality."* Sound advice, indeed. But what Jenks and Brown may not have realized was that the technology they were pioneering would one day see wide use by users unaware of any such trade-off.

In a fairly elaborate experiment that saw us measure and integrate psychometrics with intuition, preference and performance measures, we validated several concepts and gained new insights into others. We replicated, again, the distinction between display formats supporting shape understanding versus those supporting relative position (St. John et al., 2001a). Topographic views best supported the precise elevation estimation needed for Phases 2 & 4, whereas 45˚ shaded perspective views best supported extracting shape and layout for appreciating where to lay a concealed route through terrain in Phases 1 & 3.

We refined our earlier understanding of the displays necessary to support routing through terrain (St. John et al., 2001b). Like that earlier study, we found that initially determining concealed routes was done fastest with shaded relief. But unlike that earlier study, we found detailed route laying was performed best with a shaded 3-D view, as opposed to a 2-D topo view. This difference may be explained by the local nature of line of sight judgments in the earlier Antenna task compared to the global judgments required in the current routing task. Antenna required precise, local line of sight judgments from one point to just one or two other points, at most. Here, our Phase 3 task required laying an entire route concealed from as much of the terrain as possible (every point had to be concealed from as many other points as possible). This local/global explanation makes a prediction that we intend to test. 2-D superiority should gradually give way to 3-D superiority as the task requirements are smoothly changed from local focus to global concealment by, say, gradually increasing the number of points to remain concealed from.

Predictions from the Naïve Realism theory were borne out. With respect to the key Phase 3 routing task, it correctly predicted that participants would prefer and predict best performance for the high fidelity shaded, 45˚ perspective view. Yet a lower fidelity, less realistic display proved best for that task, and by a large margin. Smoothing improved routing performance, despite intuitions to the contrary. In addition, there was some Naïve Realism for the altitude reconstruction tasks. Only after they had performed the elevation reconstruction task did participants modify their assessments that the topo format was going to be most helpful. Initially, many believed they could judge altitude better from the realistic shading.

We went beyond merely documenting the Naïve Realism of naïve users, though. We began to investigate what characteristics may be associated with Naïve Realism. Although preliminary, an interesting interaction with cognitive styles and spatial ability emerged. High-spatials performed Phase 2-4 tasks better and they better calibrated their intuitions and preferences for appropriate displays after using them. Low-spatials performed Phase 2-4 tasks worse and were not as good at calibrating their intuitions and preferences for appropriate displays. Interestingly, comparable results have recently been observed in the naturalistic preference and use of graphical weather maps by Navy weather forecasters (Smallman & Hegarty, this meeting). That the same trend is emerging in disparate domains raises another interesting question. Do low-spatials perform this way because of an insensitivity to, or an inability to take advantage of, experience with displays?

In applied implications of the work, we developed a terrain appreciation display concept with potential for operational use. Past research has only used terrain simplification as an experimental control condition. Eley (1991) smoothed terrain to create control displays for testing whether memory encodes only the gist of terrain. Here, we are proposing applying terrain simplification to actual task displays. Filtering the terrain has the desirable consequence of obliterating details that can mask the detection of gross scene features necessary to get scene layout. For example, one doesn't need to see the jagged edges of a canyon to detect the canyon itself.

The final questions that this line of research has opened up are how much terrain smoothing to employ, and how and whether to put simplification under user control. Simple terrain, such as glaciated terrain, may not need much filtering, whereas more complex mountainous terrain may benefit from more. In terms of user control, we have implemented a prototype system that has a continuous smoothing dial. The idea is that users could smoothly dial down terrain complexity for global scene reorientation as they engage in complex tasks. We will investigate its utility in future work.

ACKNOWLEDGEMENTS

This work was sponsored by the Office of Naval Research.

REFERENCES

Eley, M. G. (1991). Selective encoding in the interpretation of topographic maps. *Applied Cognitive Psychology*, 5, 403-422.

Hegarty, M. & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91, 684-689.

Jenks, G.F., & Brown, D.A. (1966) Three dimensional map construction. *Science*, 154, 857-864.

Kozhevnikov, M., Hegarty, M. & Mayer, R. E. (2002). Revising the visualizer/verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20, 47-77.

Pick, H. L., Heinrichs, M. R., Montello, D. R., Smith, K., Sullivan, C. N., & Thompson, W. B. (1995). Topographic map reading. In P. A. Hancock, J. M. Flach, J. Caird, & K. J. Vicente (Eds.), *Local applications of the ecological approach to human-machine systems, Vol.2* (pp. 255-284). Hillsdale, NJ: Lawrence Erlbaum.

Savage, D.M., Wiebe, E.N., & Devine, H.A. (2004) Performance of 2D versus 3D topographic representations for different task types. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 1793-1797).

Smallman, H.S. & Hegarty, M. (2007) Expertise, spatial ability and intuition in the use of complex visual displays. *This meeting*.

Smallman, H.S., & St. John, M. (2005) Naïve Realism: Misplaced faith in the utility of realistic displays. *Erg. in Design*, 13, 6-13.

St. John, M., Cowen, M.B., Smallman, H.S., & Oonk, H.M. (2001a) The use of 2D and 3D displays for shape understanding vs. relative position tasks. *Human Factors*, 43, 79-98.

St. John, M., Smallman, H.S., Bank, T.E., & Cowen, M.B. (2001b) Tactical route planning using two-dimensional and three-dimensional views of terrain. In *Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society*. (pp. 1409-1413).